



## Efficient Incorporation of Markov Random Fields in Change Detection

Aanæs, Henrik; Nielsen, Allan Aasbjerg; Carstensen, Jens Michael; Larsen, Rasmus; Ersbøll, Bjarne Kjær

*Published in:*  
IEEE International Geoscience and remote sensing symposium

*Link to article, DOI:*  
[10.1109/IGARSS.2009.5417856](https://doi.org/10.1109/IGARSS.2009.5417856)

*Publication date:*  
2009

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Aanæs, H., Nielsen, A. A., Carstensen, J. M., Larsen, R., & Ersbøll, B. K. (2009). Efficient Incorporation of Markov Random Fields in Change Detection. In *IEEE International Geoscience and remote sensing symposium* (Vol. 3). IEEE. <https://doi.org/10.1109/IGARSS.2009.5417856>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# EFFICIENT INCORPORATION OF MARKOV RANDOM FIELDS IN CHANGE DETECTION

*H. Aanæs<sup>1</sup>, A. A. Nielsen<sup>2</sup>, J. M. Carstensen<sup>1</sup>, R. Larsen<sup>1</sup> and B. Ersbøll<sup>1</sup>*

DTU Informatics<sup>1</sup> & DTU Space<sup>2</sup>,  
Technical University of Denmark,  
DK-2800 Kgs. Lyngby, Denmark  
e-mail: haa@imm.dtu.dk

## ABSTRACT

Many change detection algorithms work by calculating the probability of change on a pixel-wise basis. This is a disadvantage since one is usually looking for regions of change, and such information is not used in pixel-wise classification – per definition. This issue becomes apparent in the face of noise, implying that the pixel-wise classifier is also noisy. There is thus a need for incorporating local homogeneity constraints into such a change detection framework. For this modelling task Markov Random Fields are suitable. Markov Random Fields have, however, previously been plagued by lack of efficient optimization methods or numerical solvers. We here address the issue of efficient incorporation of local homogeneity constraints into change detection algorithms. We do this by exploiting recent advances in graph based algorithms for Markov Random Fields. This is combined with an IR-MAD change detector, and demonstrated on real data with good results.

**Index Terms**— Change Detection, Markov Random Fields, Homogeneity Constraints, Graph Based Algorithms, IR-MAD.

## 1. INTRODUCTION

Remotely sensed Earth observation data are typically collected as reflection or emission of electromagnetic radiation from the surface of the Earth. The data are collected globally, usually in a digital image format, and cover the same geographical area after some period of time. For many Earth observation platforms data have been available for decades. Therefore one of the most important applications of this type of data is the study of spatio-temporal dynamics of the surface of the Earth related to areas such as agriculture, forestry, environment, oceanography, etc. Such studies include change detection over time in multi- and hyperspectral digital image data.

A large class of change detection methods works by calculating a statistical probability of change for each pixel, as seen in Figure 4. This has the disadvantage that in many applications one is looking for regions of change, and such infor-

mation is, by definition, not used in pixel-wise classification. To elaborate, an underlying problem is that the probability distributions describing change and no-change have considerable 'overlap', and thus there is a lot of noise in the pixel-wise change detection result, as seen in Figure 5. In particular clearly defined regions of change are hard to identify. Therefore there is a need to incorporate a notion of local homogeneity in the modelling of the problem, i.e. the solution to the change detection should be made considering the fact that we expect homogenous regions, and thus a single pixel of change in a large region of no-change is very unlikely. This is the issue which we address here.

A good way of incorporating such homogeneity constraints into the modelling of the problem is via Markov Random Fields (MRF) cf. e.g. [1, 2]. MRFs have been used successfully to solve similar problems before, and they have a very well developed and understood underlying theory. This is what we propose doing here.

Previously, the problem using MRF for such tasks has been one of optimization, or numerically finding an optimal solution. Typically simulated annealing or other stochastic optimization techniques have been used. These methods are slow and seldom give the optimal result c.f. e.g. [3]. However, it was relatively recently discovered that graph based techniques could be used to solve these problems, yielding highly efficient algorithms achieving guaranteed optimal results in the two class case, cf. e.g. [4]. We therefore employ these methods in the present work. Doing so ensures that the results achieved are a result of one's modelling of the problem, and are not influenced by stochastic artifacts due to the optimization method. This enables more rigorous experimentation with regularization parameters, e.g. in a regularization path framework [5, 6].

The use of such homogeneity constraints in change detection is ongoing work, and we aim to compare with many other probability functions, and extend to more than two classes (i.e. change vs. no-change). However, we feel that the results presented here are encouraging. We thus propose using the novel combination of the IR-MAD change detector with MRF, and to do it in an efficient way which is guaranteed to

achieve the optimum of the objective function. The obtained results are promising. Other change detection schemes based on MRF are [7, 8], which use automatic threshold selection and the expectation-maximization (EM) algorithm, [9], for unsupervised estimation of the statistics of the changed and unchanged pixels in the simple difference image.

## 2. MARKOV RANDOM FIELDS & CHANGE DETECTION

Pixel-wise change detection is based on; for each pixel,  $x_i$ , calculating the probability,  $P_{image}(x_i)$ , of change occurring in that pixel. Then if  $P_{image}(x_i) > 0.5$  (or some other threshold) the pixel is classified as 'change'. The pixel-wise probability measure,  $P_{image}(x_i)$ , used here is from the so-called IR-MAD change detector [10, 11].

The MAD method is based on the established technique of canonical correlation analysis: for the multivariate data acquired at two points in time and covering the same geographical region, we calculate the canonical variates and subtract them from each other. These orthogonal differences contain maximum information on joint change in all variables (spectral bands). The change detected in this fashion is invariant to separate linear (affine) transformations in the originally measured variables at the two points in time, such as 1) changes in gain and offset in the measuring device used to acquire the data, 2) data normalization or calibration schemes that are linear (affine) in the gray values of the original variables, or 3) orthogonal or other affine transformations, such as principal component (PC) or maximum autocorrelation factor (MAF) transformations. The IR-MAD method first calculates ordinary canonical and original MAD variates. In the following iterations we apply different weights to the observations, large weights being assigned to observations that show little change, i.e., for which the sum of squared, standardized MAD variates is small, and small weights being assigned to observations for which the sum is large. Like the original MAD method, the iterative extension is invariant to linear (affine) transformations of the original variables.

The probability measure,  $P_{image}(x_i)$ , can be combined with a homogeneity term or constraint in an MRF framework, as previously mentioned. We propose doing this by forming the following probability measure for all pixels, denoted  $x$ :

$$P(x) = \frac{1}{Z} \exp \left( \sum_i \left\{ \log(P_{image}(x_i)) + \beta \sum_{j \in \mathcal{N}(x_i)} n_{ij} \right\} \right), \quad (1)$$

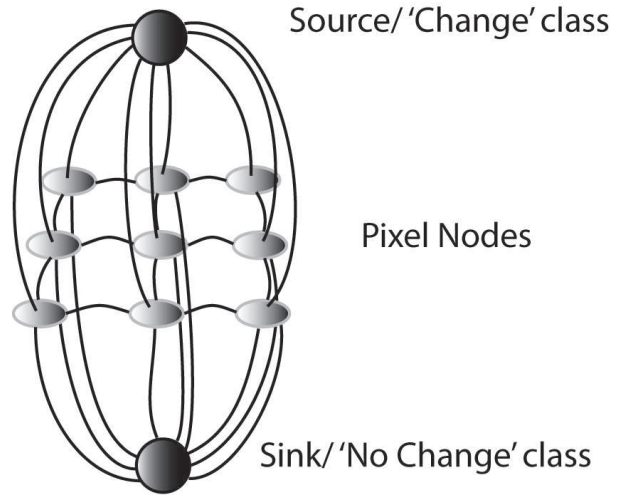
where  $Z$  is a normalization constant and  $n_{ij}$  is a function that is one if pixel  $x_i$  and  $x_j$  are assigned the same class<sup>1</sup> and zero otherwise. The neighborhood,  $\mathcal{N}(x_i)$ , is here chosen to be a standard 4-neighborhood. This is the so-called Ising model [2], which was first used for describing magnetism in iron.

<sup>1</sup>Here there are two classes; change and no-change.

The  $\beta$  parameter<sup>2</sup> determines the degree of regularization, see Figures 6 and 7.

In (1), which is an MRF, it is seen that the second term, i.e.  $\beta \sum_{j \in \mathcal{N}(x_i)} n_{ij}$ , is the homogeneity constraint. The  $n_{ij}$  are one if two neighboring pixels have the same class, and zero otherwise. It is seen that the probability of an image classification is in general increased if neighboring pixels are assigned the same class. This is again equivalent to a homogeneity term.

## 3. GRAPH BASED ALGORITHMS



**Fig. 1.** Illustration of a graph, representing an MRF. A MAP solution to the MRF can be found by calculating the minimal cut in this graph. The edges (black lines) are associated with a cost. After a minimal cut is found, pixel nodes in the same 'cut' as the source node, will be classified as 'change' pixels.

An efficient way of solving MRF such as in (1), is via graph cut algorithms, cf. e.g. [4]. If there are only two classes<sup>3</sup> the problem can be cast as a graph problem as illustrated in Figure 1. This graph contains a node for each pixel as well as two extra nodes — the source and the sink. The edges of this graph contain the costs of (1). The costs associated with the edges between the pixel nodes are  $\beta$ , thus capturing the second term of (1). The first term of (1), associated with  $P_{image}(x_i)$  is captured by the edges from the pixel nodes to the source and the sink.

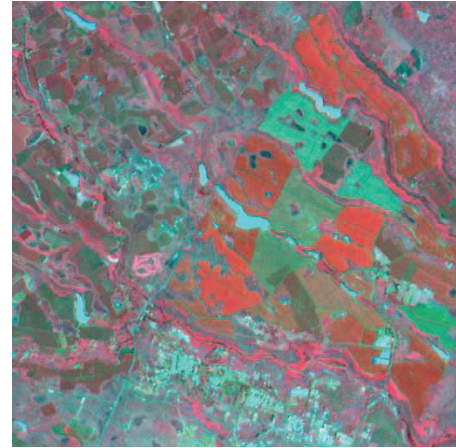
It has been shown, [4], that finding the minimum cut of the graph illustrated in Figure 1, and described above, correspond to minimizing (1), thus solving the problem. A minimum cut is the partition of the graph (with the source and sink

<sup>2</sup> $\beta$  should be nonnegative.

<sup>3</sup>In case of more than two classes this method can be adapted via a so called  $\alpha$ -expansion.



**Fig. 2.** Image of Thika from 1987.



**Fig. 3.** Image of Thika from 1989.

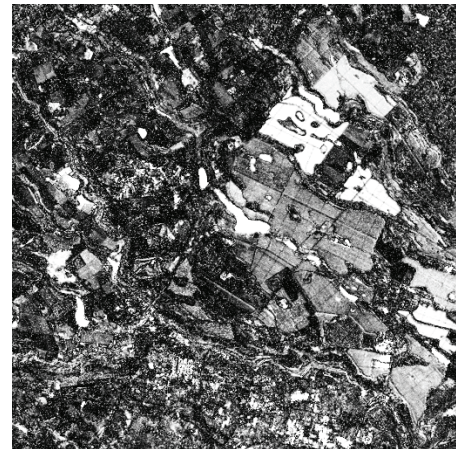
in different partitions) such that the cost of the edges severed is minimal. Efficient algorithms exist for solving such minimal cost problems, making graph cuts an efficient method for solving this type of MRF.

#### 4. EXPERIMENTAL RESULTS

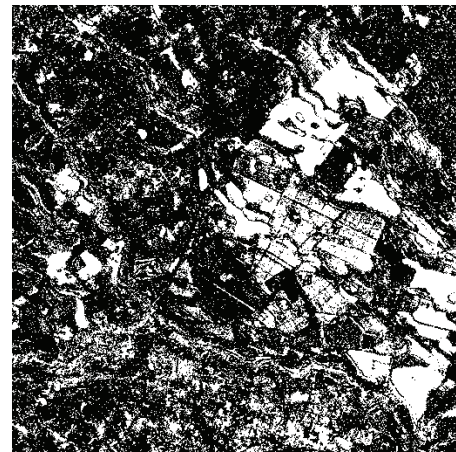
To illustrate and validate our proposed approach, we applied it to two images of Thika, which lies to the immediate north of Nairobi, Kenya — see Figures 2 and 3. The images are 512 by 512 20 m pixels SPOT HRV multispectral data acquired on 5 Feb 1987 and 12 Feb 1989 respectively. The geographical region covered is an agricultural area with large scale pineapple (center and right), small scale coffee (left). The color encoding shows the photo-infrared band as red, red as green, and green as blue. The pixel-wise probability measure of change is illustrated in Figure 4, calculated via the IR-MAD change detector [10, 11]. The classical thresholded result is depicted in Figure 5, and the result of the proposed method is seen in Figures 6 and 7, with varying  $\beta$  (i.e. strength of the homogeneity constraint). From these results it is seen that the proposed method works as expected, giving homogeneous and meaningful results. The effect of  $\beta$  is also illustrated in Figures 6 and 7, where it is clearly seen that  $\beta$  controls how homogenous the result should be.

#### 5. CONCLUSION

We have proposed using the novel combination of the IR-MAD change detector with MRF, and do it in an efficient way which is guaranteed to achieve the optimum of the objective function. This has been demonstrated to give good results, implying that MRFs are useful for this type of task.

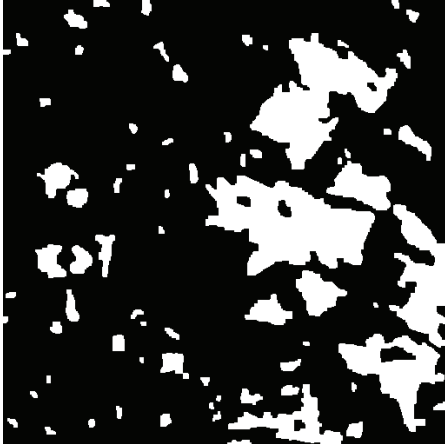


**Fig. 4.** Probability of change between Figures 2 and 3, without a homogeneity constraint.

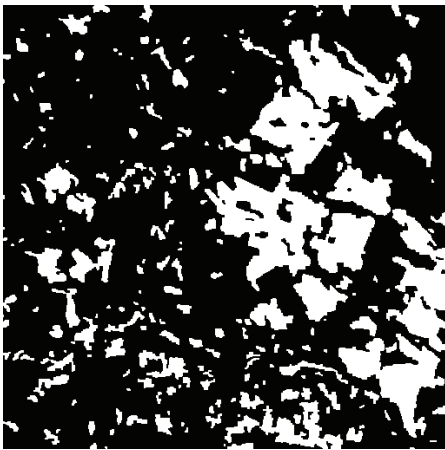


**Fig. 5.** The image of Figure 4, with a threshold of 0.5. This represents a classical change detection scheme.





**Fig. 6.** The proposed approach for change detection, with  $\beta = 1$ .



**Fig. 7.** The proposed approach for change detection, with  $\beta = 0.5$ .

## 6. REFERENCES

- [1] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (Applications of Mathematics)*, Springer Verlag, 2003.
- [2] S.Z. Li, *Markov Random Field Modeling in Image Analysis*, Springer Verlag, 2nd edition, 2007.
- [3] D. M. Greig, B. T. Porteous, and A. H. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 51, no. 2, pp. 271–279, 1989.
- [4] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [5] K. Sjöstrand, M.S. Hansen, H.B. Larsson, and R. Larsen, "A path algorithm for the support vector domain description and its application to medical imaging," *Medical Image Analysis*, vol. 11, no. 5, pp. 417–428, 2007.
- [6] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net.," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301, 2005.
- [7] L. Bruzzone and D.F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1171–1182, 2000.
- [8] L. Bruzzone and D.F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 452–466, 2002.
- [9] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [10] A.A. Nielsen, "The regularized iteratively reweighted mad method for change detection in multi- and hyperspectral data," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 463–478, 2007.
- [11] M.J. Canty and A.A. Nielsen, "Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted mad transformation," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 1025–1036, 2008.
- [12] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.